

\* NOTICES \*

JPO and NCIP are not responsible for any damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.\*\*\* shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

---

CLAIMS

---

[Utility model registration claim]

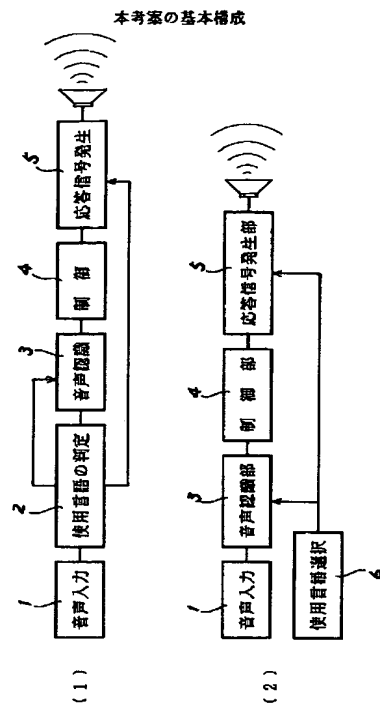
[Claim 1] In the voice recognition unit which will output the talk back signal corresponding to this sound signal, and will answer with voice if the signal of the voice which the speaker uttered inputs The language judging section used which judges the language from the sound signal inputted as the voice input section (1) at least (2), The speech recognition section (3) which distinguishes the contents of the inputted sound signal, and the control section which controls a device according to a speech recognition result (4), It has the reply signal generating section for an audio response (5). The speech recognition section (3) Two or more languages can be recognized from the inputted sound signal. Moreover, the reply signal generating section (5) The audio response by two or more languages has become possible, and the judgment signal outputted from the language judging section (2) used inputs into the speech recognition section (3). The voice recognition unit characterized by for the judgment signal which recognition object language was selected and was outputted from the language judging section (2) used inputting into the reply signal generating section (5), and being constituted so that audio response language may be selected.

[Claim 2] In the voice recognition unit which will output the talk back signal corresponding to this sound signal, and will answer with voice if the signal of the voice which the speaker uttered inputs The speech recognition section which distinguishes the contents of the sound signal inputted as the voice input section (1) at least (3), It has the control section (4) which controls a device according to a speech recognition result, the reply signal generating section for an audio response (5), and the setting section (6) of the language used. The speech recognition section (3) Two or more languages can be recognized from the inputted sound signal. Moreover, the reply signal generating section (5) The audio response by two or more languages has become possible. Either [ at least ] said speech recognition section (3) or said reply signal generating section (5) The voice recognition unit characterized by being constituted according to the selection signal outputted from the language setting section (6) used so that recognition object language or audio response language may be chosen.

---

[Translation done.]

Drawing selection Representative drawing



[Translation done.]

(19)日本国特許庁(JP)

(12) 公開実用新案公報(U)

(11)実用新案出願公開番号

実開平6-73800

(43)公開日 平成6年(1994)10月18日

(51)Int.Cl. <sup>5</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	5 6 1 D	9379-5H		
	5 5 1 C	9379-5H		
H 0 4 Q 9/00	3 1 1 Q	7170-5K		

審査請求 未請求 請求項の数 2 O L (全 7 頁)

(21)出願番号 実願平5-15272

(22)出願日 平成5年(1993)3月30日

(71)出願人 000237592

富士通テン株式会社

兵庫県神戸市兵庫区御所通1丁目2番28号

(72)考案者 藤本 博之

兵庫県神戸市兵庫区御所通1丁目2番28号

富士通テン株式会社内

(72)考案者 佐古 和也

兵庫県神戸市兵庫区御所通1丁目2番28号

富士通テン株式会社内

(72)考案者 藤本 昇治

兵庫県神戸市兵庫区御所通1丁目2番28号

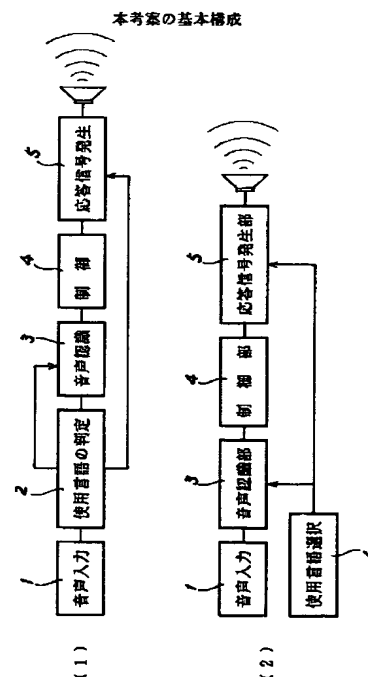
富士通テン株式会社内

(54)【考案の名称】 音声認識装置

(57)【要約】

【目的】話者の発声に対して音声で応答を行なう音声認識装置に関し、予め複数カ国語で音声認識および音声応答可能とすることで使い勝手を良くし、かつ汎用性の高い安価な装置を実現することを目的とする。

【構成】少なくとも、音声入力部1と、入力して来た音声信号からその言語を判定する使用言語判定部2と、入力された音声信号の内容を判別する複数カ国語対応の音声認識部3と、音声認識結果に従って機器を制御する制御部4と、音声応答のための複数カ国語対応の応答信号発生部5とを有し、使用言語判定部2から出力した判定信号が音声認識部3に入力して、認識対象言語が選定され、かつ使用言語判定部2から出力した判定信号が応答信号発生部5に入力して、音声応答言語が選定されるような構成とする。



## 【実用新案登録請求の範囲】

【請求項 1】 話者の発声した音声の信号が入力すると、該音声信号に対応するトークバック信号を出力し、音声で応答する音声認識装置において、少なくとも、音声入力部（１）と、入力して来た音声信号からその言語を判定する使用言語判定部（２）と、入力された音声信号の内容を判別する音声認識部（３）と、音声認識結果に従って機器を制御する制御部（４）と、音声応答のための応答信号発生部（５）とを有し、音声認識部（３）は、入力した音声信号から複数カ国語を認識でき、また応答信号発生部（５）は、複数カ国語による音声応答が可能となっており、使用言語判定部（２）から出力した判定信号が音声認識部（３）に入力して、認識対象言語が選定され、かつ使用言語判定部（２）から出力した判定信号が応答信号発生部（５）に入力して、音声応答言語が選定されるように構成されていることを特徴とする音声認識装置。

【請求項 2】 話者の発声した音声の信号が入力すると、該音声信号に対応するトークバック信号を出力し、音声で応答する音声認識装置において、少なくとも、音声入力部（１）と、入力された音声信号の内容を判別する音声認識部（３）と、音声認識結果に従って機器を制御する制御部（４）と、音声応答のための応答信号発生部（５）と、使用言語の設定部（６）とを有し、音声認識部（３）は、入力した音声信号から複数カ国語

を認識でき、また応答信号発生部（５）は、複数カ国語による音声応答が可能となっており、前記音声認識部（３）および前記応答信号発生部（５）の少なくとも一方は、使用言語設定部（６）から出力された選択信号に応じて、認識対象言語あるいは音声応答言語が選択されるように構成されていることを特徴とする音声認識装置。

## 【図面の簡単な説明】

【図 1】 本考案による音声認識装置の基本原理を説明するブロック図である。

【図 2】 請求項 1 の音声認識装置を自動車に実施した例を示すブロック図である。

【図 3】 図 2 の実施例装置における処理手順を例示するフローチャートである。

【図 4】 請求項 2 の考案における処理手順を例示するフローチャートである。

【図 5】 従来の音声認識装置における応答動作を示すフローチャートである。

## 【符号の説明】

1 音声入力部

2 使用言語判定部

3 音声認識部

4 制御部

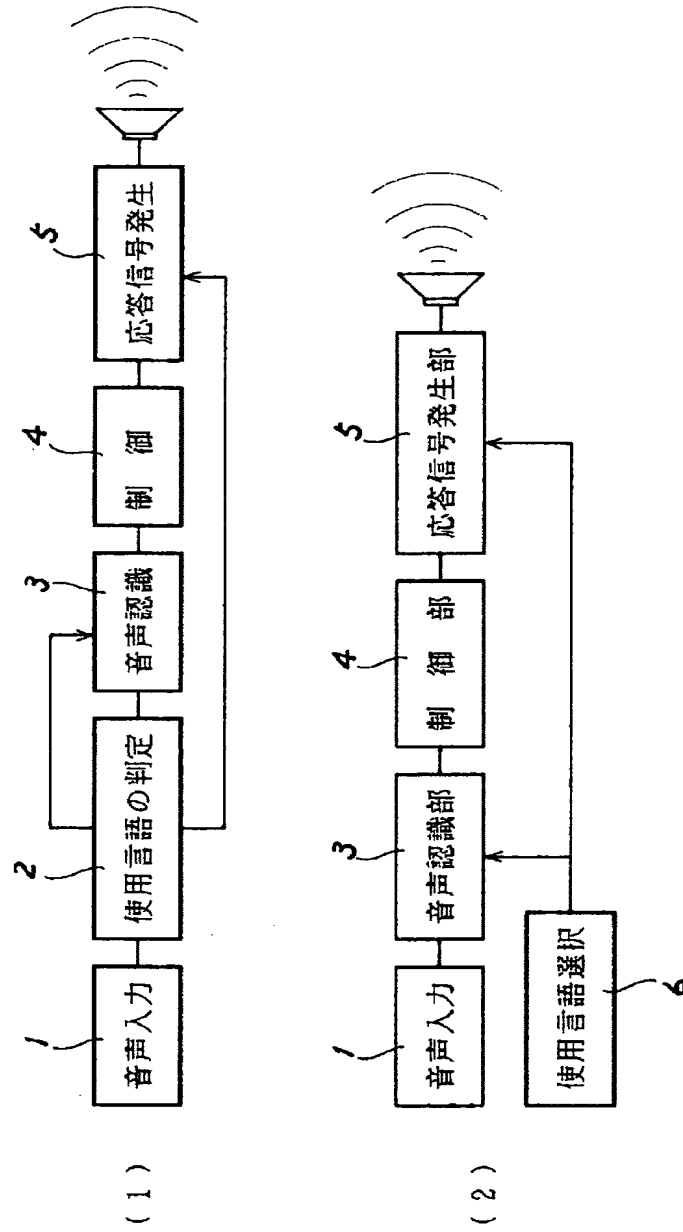
5 応答信号発生部

6 使用言語設定部

A1～An 自動車における各種装備

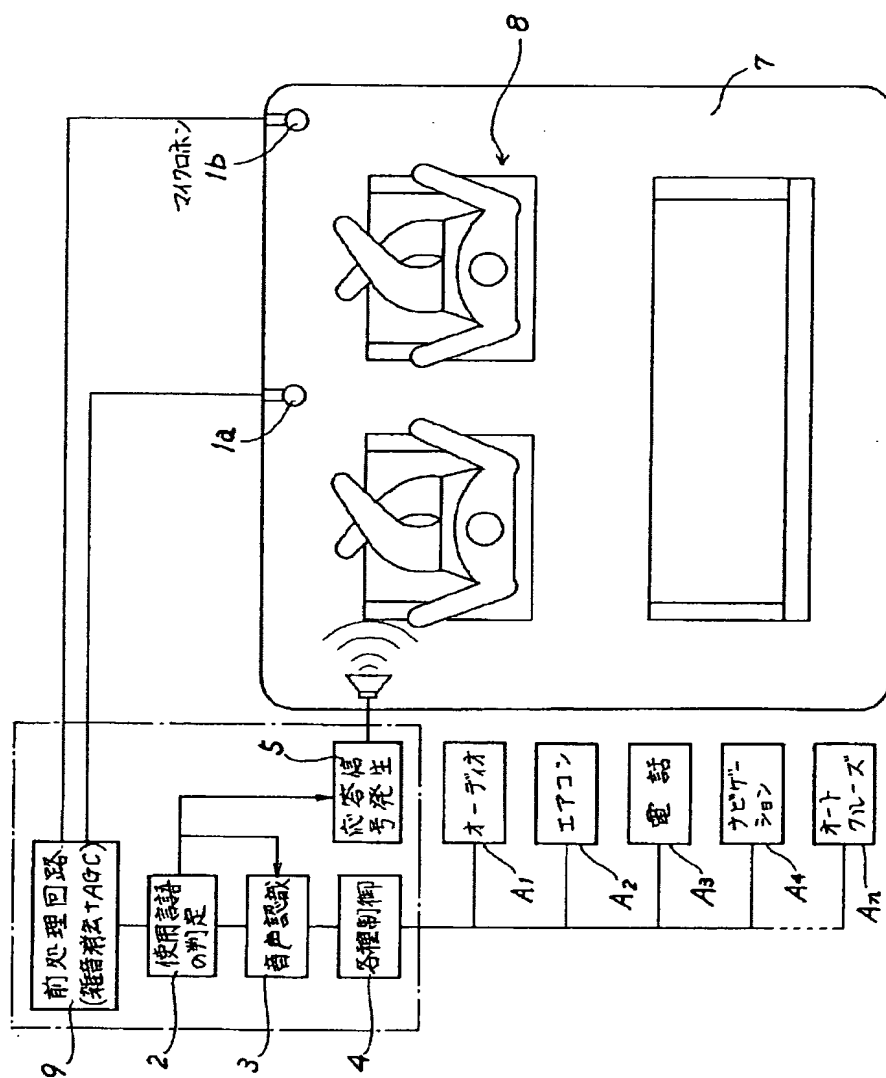
【図1】

本考案の基本構成



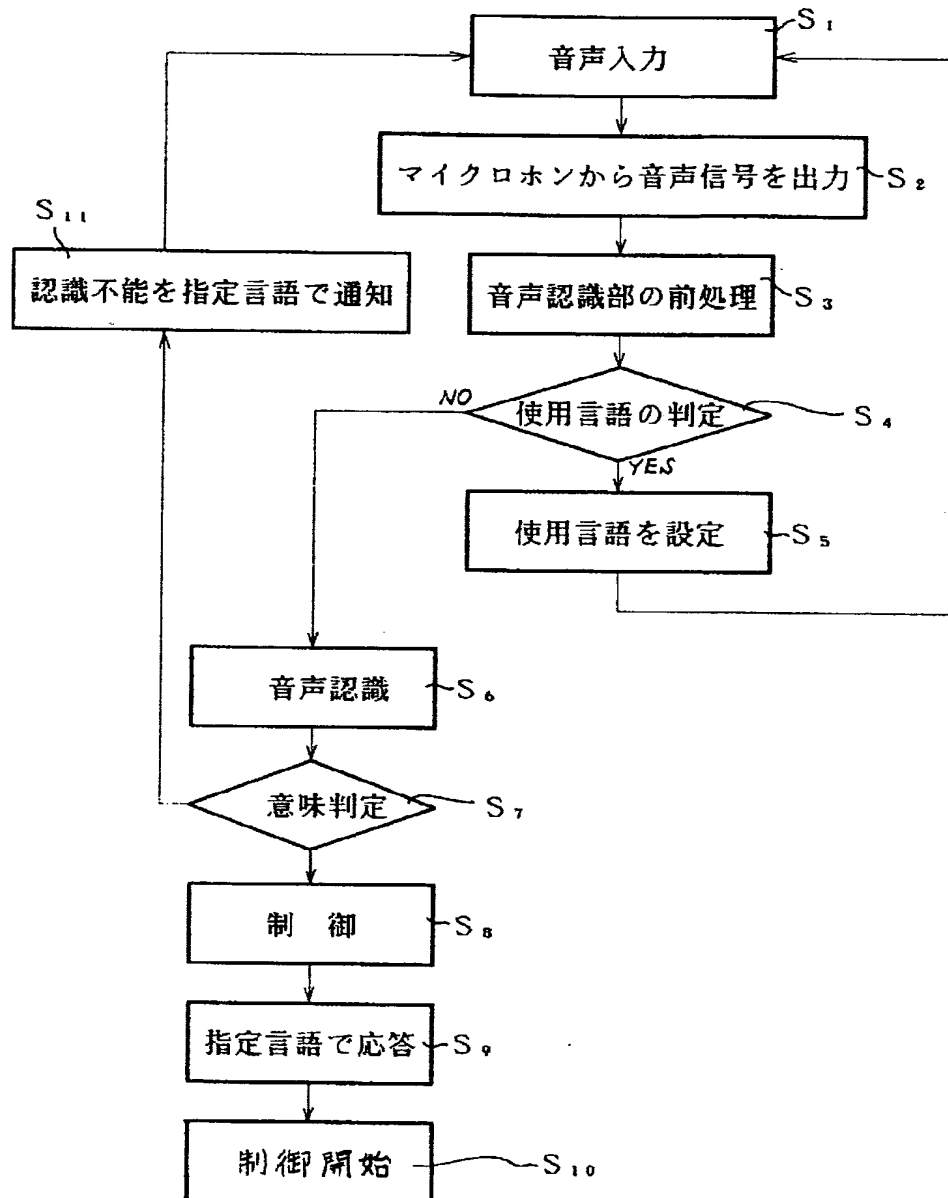
【図2】

自動車に実施した例



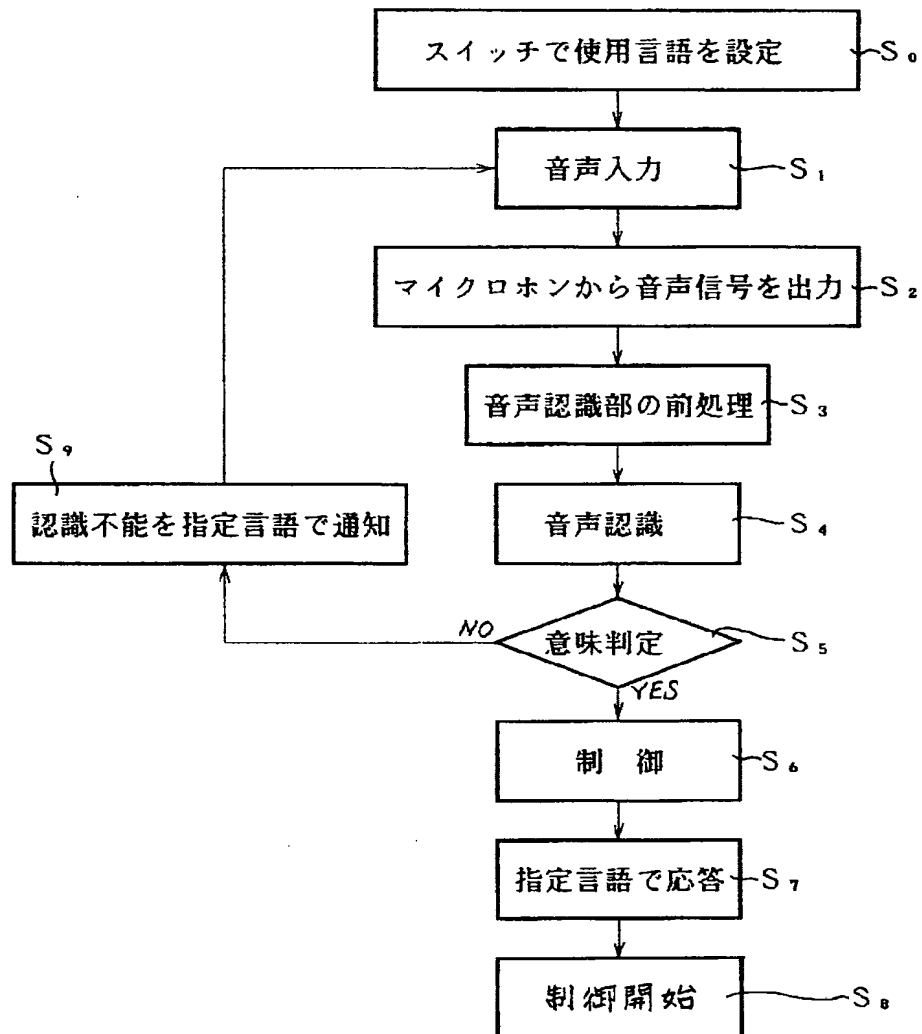
【図3】

## 請求項1の実施例



【図4】

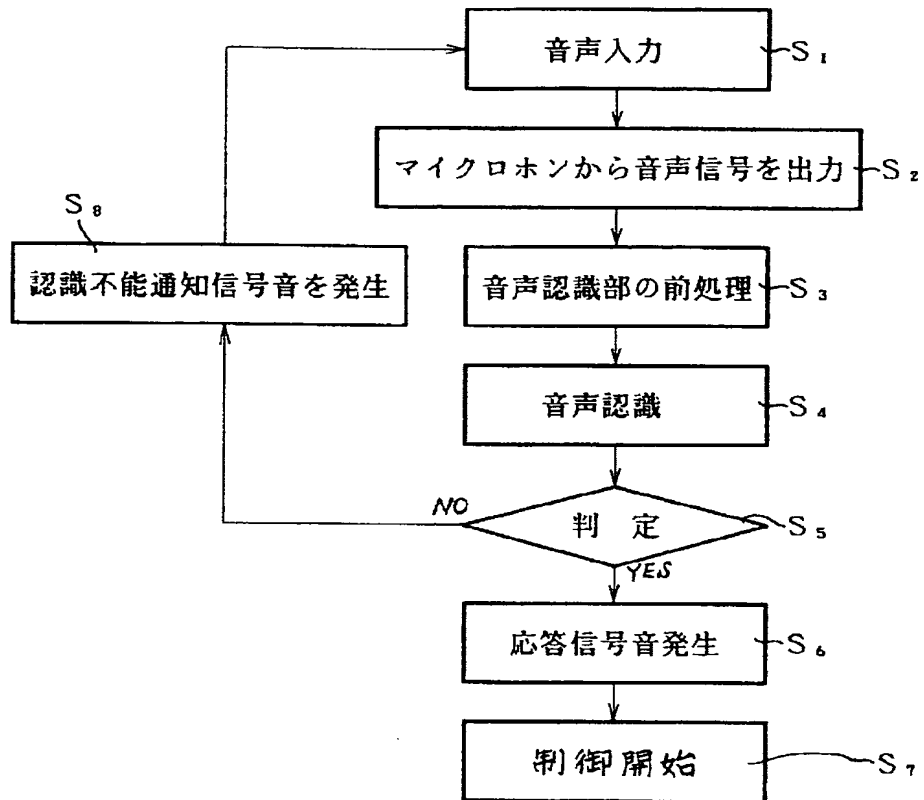
## 請求項2の実施例





【図5】

## 従来の音声認識装置



**【考案の詳細な説明】****【0001】****【産業上の利用分野】**

本考案は、話者の発声に対して音声で応答を行なう音声認識装置に関する。自動車などの運転者が運転中に、エアコンディショナやオーディオ装置、カーナビゲーション、自動車電話、オートクルーズなどの機器を手動操作すると、安全運転に支障を来し、事故を引き起こすおそれがある。そこで、これらの機器の操作を音声で指示し、また機器側では、指示された操作に対して音声で応答して、話者が確認できるようにすることが、操作性向上のために望ましい。

**【0002】****【従来技術】**

自動車用に限らず、音声認識装置は、話者の音声を認識し、さらに認識した内容に従って機器を操作するようになっている。そして、結果を話者に対して、機器側から信号音を発して応答する。

**【0003】**

図5は従来の音声認識装置における応答動作を示すフローチャートである。オーディオ機器を操作する場合であれば、まず、ステップS1において、例えば「プレイ」と発声して音声入力し、再生動作を指示すると、ステップS2において、マイクロホンを通じて音声信号が入力され、ステップS3で音声認識に適するように前処理が行なわれる。

**【0004】**

次いで、ステップS4に示すように、音声認識部において、音声認識が行なわれ、ステップS5において認識結果の判定が行なわれる。「プレイ」という言葉を認識できた場合は、ステップS6において、認識できたことをブザー音などで応答して話者に知らせた後、ステップS7において、プレイ動作を開始する。

**【0005】**

これに対し、話者の言葉を認識できない場合は、ステップS8において、その旨を知らせる信号音を出力して、話者に対し、再度音声入力するよう促す。

**【0006】**

**【考案が解決しようとする課題】**

しかしながら、このようなブザー音などによる応答では、音声認識できたか、できないかの二通り程度しか表現できない。自動車などの場合は、多数の機器が装備されており、多数のスイッチを操作する必要があるので、スイッチ操作によって行なわれる動作の内容も多種多様である。したがって、応答もブザー音ではなく、操作指示の内容に応じて具体的に音声で応答できるのが望ましい。

**【0007】**

例えば、話者が「プレイ」と発声した場合、その指示音声を確認し、しかもプレイ可能であれば、例えば「カセットテープをプレイします」のように、音声で応答し、音声認識できない場合は、「もう一度お話し下さい」のように、音声応答（トークバック）できると、対話式に、より正確に機器操作を行なうことができ、操作性が向上する。

**【0008】**

また、エアコンディショナの場合は、温度や風量などの状態を、オートクルーズの場合は、現在の速度などの状態を、音声応答で話者に知らせることができると、操作がより簡便となる。

**【0009】**

ところで、音声認識は通常、日本語で行なわれるが、輸出車の場合は、相手国語で音声認識可能にしておく必要がある。また、本考案のように音声で応答する場合も、輸出用の車両については、相手国語で応答する必要がある。

**【0010】**

しかしながら、音声認識やトークバックを、輸出車に関してのみ、英語などの相手国語で行なえるようにするには、英語を音声認識可能とし、しかもトークバックも英語で行なえる認識装置を別に開発する必要がある、コスト高となるという問題がある。

**【0011】**

一方、国際化が進展すると、レンタルカーなどを想定したとき、同一車両を日本人が運転することもあれば外国人が運転することもある。このような場合、日本語のみ又は英語のみしか音声認識やトークバックできないとなると、所期の目

的を達成できず、不便である。

【0012】

本考案の技術的課題は、このような問題に着目し、予め複数カ国語で音声認識および音声応答可能とすることで使い勝手を良くし、かつ汎用性の高い安価な装置を実現することにある。

【0013】

【課題を解決するための手段】

図1は本考案による音声認識装置の基本構成を説明するブロック図である。請求項1の音声認識装置は、話者の発声した音声の信号が入力すると、該音声信号に対応するトークバック信号を出力し、音声で応答する音声認識装置において、図1(1)のように、少なくとも、音声入力部1と、入力して来た音声信号からその言語を判定する使用言語判定部2と、入力された音声信号の内容を判別する音声認識部3と、音声認識結果に従って機器を制御する制御部4と、音声応答のための応答信号発生部5とを有している。

【0014】

前記の音声認識部3は、入力した音声信号から複数カ国語を認識でき、また応答信号発生部5は、複数カ国語で音声応答できる。そして、使用言語判定部2から出力された判定信号が音声認識部3に入力して、認識対象言語が選定され、かつ使用言語判定部2から出力された判定信号が応答信号発生部5に入力して、音声応答言語が選定されるように構成されている。

【0015】

請求項2は、話者の発声した音声の信号が入力すると、該音声信号に対応するトークバック信号を出力し、音声で応答する音声認識装置において、図1(2)のように、少なくとも、音声入力部1と、入力された音声信号の内容を判別する音声認識部3と、音声認識結果に従って機器を制御する制御部4と、音声応答のための応答信号発生部5と、使用言語の設定部6とを有している。

【0016】

音声認識部3は、入力した音声信号から複数カ国語を認識でき、また応答信号発生部5は、複数カ国語による音声応答が可能となっている。そして、前記音声

認識部3および前記応答信号発生部5の少なくとも一方は、使用言語設定部6から出力された選択信号に応じて、認識対象言語あるいは音声応答言語が選択されるように構成されている。

#### 【0017】

##### 【作用】

本考案の音声認識部3は、それぞれの言語ごとに音声認識機能を有していて、使用言語判定部2からの判定信号によって、判定された言語の音声認識機能が作動する構成になっている。したがって、入力した音声信号から複数カ国語を認識できる。

#### 【0018】

また、応答信号発生部5は、それぞれの言語ごとに、操作命令に応じた音声応答機能を有しているため、使用言語判定部2において、入力して来た音声信号からその言語が自動的に判定されて、判定信号が音声認識部3と応答信号発生部5に入力し、使用言語が指定されると、音声認識部3では指定された言語のみを選択して音声認識し、応答信号発生部5では、指定された言語を選択して音声応答が行なわれる。

#### 【0019】

このように、音声入力部1から音声が入力するだけで、その言語が判定され、その言語で音声認識や音声応答が行なわれるので、複数カ国語で音声認識および音声応答でき、輸出用の装置や複数カ国の人種が使用する場合に便利である。また、輸出先の国語専用の装置とする必要がなく、汎用性が高いので、安価に実現できる。

#### 【0020】

なお、使用言語判定部2においては、初期状態で入力してきた最初の音声信号で言語の種別を判定して、以後はその言語のみを有効とする手法と、音声信号が入力してくるたびにその言語の種別を判定し、判定信号を音声認識部3および応答信号発生部5に入力して、使用言語を指定する手法とがあるが、本考案は両方を含むものとする。

#### 【0021】

請求項2における音声認識部3も、入力した音声信号から複数カ国語を認識でき、また応答信号発生部5は、複数カ国語で音声応答が可能となっており、使用言語設定部6から出力した設定信号が音声認識部3に入力して、認識対象言語が選択され、あるいは使用言語設定部6から出力した設定信号が応答信号発生部5に入力して、音声応答言語が選択される構成になっている。

#### 【0022】

したがって、予め人為的に使用言語を選択する操作が必要ではあるが、複数カ国語に対応できるので、請求項1の場合と同様な効果を奏するほか、一旦使用言語を設定すれば、使用言語を変更しない限り、操作を行なう必要がないので、使用言語の変更の頻度が少ない装置に有利であり、また構成が簡素となるため、より安価に実現できる。

#### 【0023】

##### 【実施例】

次に本考案による音声認識装置が實際上どのように具体化されるかを実施例で説明する。図2は請求項1の音声認識装置を自動車に実施した例を示すブロック図である。自動車7のフロントシート8の前方には、運転席側および助手席側の双方に、音声入力手段として、マイクロホン1a、1bが配設されている。

#### 【0024】

そして、マイクロホン1a、1bで検出された音声信号は、雑音除去などを行なう前処理回路9を経て、使用言語の判定部2に入力され、フロントシート8における話者の話した言語の種別が判別される。使用言語判定部2において、話者の使用言語が判別されると、その判別信号が音声認識部3と応答信号発生部5に入力され、判定された言語のみが有効となる。

#### 【0025】

この自動車には、装備として、オーディオ装置A1、エアコンディショナA2、自動車電話A3、ナビゲーション装置A4およびオートクルーズAnなどが搭載されている。

#### 【0026】

この実施例は、初期状態で入力してきた最初の音声信号で言語の種別を判定し

て、以後はその言語のみを有効とする構成になっているものとする。いま、フロントシート8の話者が、日本語で機器の操作を行なう場合であれば、エンジンを始動させて初期状態において、まず「ニホンゴ」と発声すると、マイクロホン1aまたは1bから入力した音声信号は、前処理回路9を経て、使用言語判定部2に入力し、使用言語の種別が「日本語」とであると判定される。

【0027】

そして、この判定信号によって、音声認識部3および応答信号発生部5では、日本語のみが有効とされ、音声認識部3に登録されている日本語の単語ないし文章と、入力してきた単語ないし文章とが比較照合される。

【0028】

例えば、オーディオ装置A1のカセットテープを再生する場合であれば、話者がまず「テープ・プレイ」と発声すると、マイクロホン1aまたは1bから入力した音声信号は、前処理回路9および使用言語判定部2を経て、音声認識部3に入力し、音声認識部3に登録されている日本語の辞書と比較照合される。

【0029】

英語など、他の国語の辞書も登録されているが、他の国語が登録されているテーブルは選択されず、使用言語判定部2から入力した判定信号によって、日本語が登録されているテーブルのみが選択され、予め登録されている用語と比較照合される。

【0030】

その結果、カセットテープの再生命令であることが認識され、認識信号にしたがって制御部4において、カセットが挿入されていて再生が可能か確認が行なわれる。そして、再生が可能であれば、応答信号発生部5から出力する音声信号によって、例えば「テープをプレイします」と、スピーカから日本語で応答が行なわれた後、テープの再生が開始する。カセットが挿入されていない場合は、「カセットが挿入されていません。カセットを挿入して下さい」と、音声応答が行なわれる。

【0031】

一方、プレイ状態において、例えば「テープ・ストップ」と発声すると、前記

のプレイ指示の場合と同様にして、制御部 4 で停止が可能か確認が行なわれた後、応答信号発生部 5 から「テープを停止します」等の応答信号を出力してスピーカを鳴動させた後に、再生動作が停止する。

#### 【0032】

さらに、カセットテープを取り出したい場合は、「テープ・イジェクト」と発声すれば、前記の場合と同じ原理で「テープをイジェクトします」等の応答音を発生した後に、カセットテープがイジェクトされる。

#### 【0033】

フロントシート 8 の話者が、英語でハンドフリー操作を行なう場合であれば、エンジン始動後に、例えば「イングリッシュ」と発声すると、マイクロホン 1 a または 1 b から入力した音声信号が前処理回路 9 を経て、使用言語判定部 2 に入力し、使用言語の種別が「英語」と判定される。この判定信号により、音声認識部 3 および応答信号発生部 5 では、英語のみが有効とされ、音声認識部 3 に登録されている英語の辞書と、入力してきた英語の音声信号とが比較照合される。したがって、以後は英語によって、搭載機器をハンドフリー操作できる。

#### 【0034】

図 3 は、図 2 の実施例の装置における処理手順を示すフローチャートである。いま、エンジンを始動させた初期状態において、ステップ S1 で、図 2 のマイクロホン 1 a または 1 b に向かって、「ニホンゴ」と発声し音声入力すると、ステップ S2 でマイクロホン 1 a または 1 b から音声信号が出力し、ステップ S3 において図 2 の前処理回路 9 により雑音除去などが行なわれる。

#### 【0035】

そして、ステップ S 4 において、本考案による言語判定部 2 に音声信号が入力され、入力された用語が言語判定用語であると、言語の種別が判定されて、ステップ S5 で音声認識装置 3 および応答信号発生部 5 に対して、日本語が設定され、使用言語の指定が完了する。

#### 【0036】

ついで、ステップ S1 において、例えば「テープ・プレイ」などの操作命令が音声入力されると、ステップ S2 でマイクロホン 1 a または 1 b から音声信号が出力



し、ステップS3において雑音除去などの前処理が行なわれる。そして、次のステップS4で、入力された用語は言語判定用語（“ニホンゴ”または“イングリッシュ”）ではないと判定される。すなわち、操作指令の用語であるので、ステップS6で、音声認識部3により、日本語の辞書テーブルが選択されて、入力した日本語の意味が認識される。

#### 【0037】

ステップS7における意味判定の結果、正常に認識された場合であれば、ステップS8において制御部4によりカセットテープの再生が可能か確認が行なわれた後、ステップS9において、「テープを再生します」と音声応答が行なわれ、ステップS10において、カセットテープの再生動作が開始される。

#### 【0038】

ステップS6における音声認識が不可能であると、ステップS7で判定された場合は、ステップS11において、応答信号発生部5から「もう一度入力して下さい」との音声信号を出力してスピーカを駆動し、音声操作のリトライを促す。

#### 【0039】

なお、エンジンを始動させて初期状態において、ステップS1で「イングリッシュ」と発声し音声入力すると、ステップS4において「英語」と判定されて、図2における音声認識部3および応答信号発生部5には、英語が設定されるので、以後の音声入力では、英語のみで音声認識および音声応答が行なわれる。

#### 【0040】

この実施例では、エンジンを始動した初期状態において、「ニホンゴ」または「イングリッシュ」と発声して、最初に、以後の使用言語を指定するようになっている。したがって、エンジンを停止しない限り、使用言語の変更はできない。

#### 【0041】

これに対し、図3の判定ステップS4を省き、入力された用語のすべてについて、図2の使用言語判定部2における判定を行ない、音声認識部3および応答信号発生部5に言語判定信号を入力する構成にすれば、言語が入力される度に言語判定を行ない、音声認識部3において、判定された言語の登録テーブル中の用語と比較照合し、かつ応答信号発生部5からは判定された言語で音声応答を行なう

ことができる。この構成によると、使用言語を変更するたびに、エンジンを停止し再度始動して言語の設定をやり直す必要はなく、操作性が向上する。

#### 【0042】

しかしながら、「テープ・プレイ」や「テープ・イジェクト」などの操作命令語のように、日本語と英語で共通している用語を用いると、日本語で操作命令したのに英語で音声応答が行なわれるとか、逆に英語で操作命令したのに日本語で音声応答が行なわれる、といった事態が発生し、混乱を来す。これを防止するには、単語で命令しないで、「テープをプレイしなさい」や「テープをイジェクトしなさい」などのように、文章で命令する必要がある、また使用言語判定部2は、単語の特徴のみでなく、日本語と英語の構文の違いなどを利用して言語判定する必要がある。

#### 【0043】

以上は、請求項1の思想によって、入力音声から自動的に言語判定を行なう例であるのに対し、請求項2のように人為的に使用言語を設定する場合は、図4のように、ステップS0において、スイッチなどで使用言語を選択すると、図1(2)における使用言語の設定部6から設定信号が発生して、音声認識部3および応答信号発生部5に入力し、以後の使用言語が選択される。

#### 【0044】

したがって、ステップS0において、例えば英語が設定されたとすると、以後は英語による操作指令のみが有効となり、他の言語を用いる場合は、ステップS0において設定しなおす必要がある。この実施例では、言語設定スイッチを手動操作する必要はあるが、一旦設定しておくと、使用言語の変更がない限り何ら操作を行なう必要がない。言語設定は、手動操作スイッチに限らず、前記のように例えば音声で設定するようにしてもよい。

#### 【0045】

なお、音声認識部3の認識対象言語と応答信号発生部5の音声応答言語のいずれか一方のみを選択可能（他方は所定言語で行なう）としてもよく、また両者の言語を別個に独立して設定できるようにしてもよい。

#### 【0046】

以上の実施例は、日本語と英語の二カ国語を使用可能とした例であるが、これのみに限定されないことはもちろんであり、また三カ国語以上を選択使用可能とすることもできる。

【0047】

音声による操作対象は、図2における機器のほか、例えばワイパーなどのような各種のスイッチ類の操作を、可能な限り本考案による音声認識装置で行なうことができる。しかも、自動車内における機器の操作に限らず、他の機器の操作にも適用できる。

【0048】

【考案の効果】

請求項1によれば、音声入力するだけで、その言語が判定され、その言語で音声認識や音声応答が行なわれるので、複数カ国語で音声認識および音声応答でき、輸出用の装置や複数カ国の人種が使用する場合に便利である。また、輸出先の国語専用の装置とする必要がないので、汎用性が高く、安価に実現できる。

【0049】

請求項2のように、予め人為的にスイッチ操作などで使用言語を選択する構成の場合は、一旦使用言語を設定すれば、使用言語を変更しない限り、操作を要しないので、使用言語の変更の頻度が少ない装置に有利であり、構成が簡素なためより安価に実現できる。